Chapter 3

Vector spaces

In this Chapter, we provide an abstract framework which encompasses what we have seen on $\mathbb{R}^n$ and for $M_{mn}(\mathbb{R})$.

3.1 Abstract definition

Before introducing the abstract notion of a vector space, let us make the following observation. For any $X, Y, Z \in \mathbb{R}^n$, for any $\mathcal{A}, \mathcal{B}, \mathcal{C} \in M_{mn}(\mathbb{R})$ and for any $\lambda, \mu \in \mathbb{R}$ one has

(i) $(X + Y) + Z = X + (Y + Z)$,

(ii) $X + Y = Y + X$,

(iii) $X + \mathbf{0} = \mathbf{0} + X = X$,

(iv) $X - X = \mathbf{0}$,

(v) $1X = X$,

(vi) $\lambda(X + Y) = \lambda X + \lambda Y$,

(vii) $(\lambda + \mu)X = \lambda X + \mu X$,

(viii) $(\lambda\mu)X = \lambda(\mu X)$.

(i) $(\mathcal{A} + \mathcal{B}) + \mathcal{C} = \mathcal{A} + (\mathcal{B} + \mathcal{C})$,

(ii) $\mathcal{A} + \mathcal{B} = \mathcal{B} + \mathcal{A}$,

(iii) $\mathcal{A} + \mathcal{O} = \mathcal{O} + \mathcal{A} = \mathcal{A}$,

(iv) $\mathcal{A} - \mathcal{A} = \mathcal{O}$,

(v) $1\mathcal{A} = \mathcal{A}$,

(vi) $\lambda(\mathcal{A} + \mathcal{B}) = \lambda\mathcal{A} + \lambda\mathcal{B}$,

(vii) $(\lambda + \mu)\mathcal{A} = \lambda\mathcal{A} + \mu\mathcal{A}$,

(viii) $(\lambda\mu)\mathcal{A} = \lambda(\mu\mathcal{A})$.

Note that these properties are borrowed from Chapter 1 and 2 respectively. Another example which would satisfy the same properties is provided by the set of real functions defined on $\mathbb{R}$, together with the addition of such functions and with the multiplication by a scalar. In this case, the element $\mathbf{0}$ (or $\mathcal{O}$) is simply the function which is equal to 0 at any point of $\mathbb{R}$.

Question: Can one give an abstract definition for these rules ?

In the first definition, we give a more general framework in which λ and μ live. You can always think about $\mathbb{R}$ as the main example for the following definition.

Definition 3.1.1. A field $(\mathbb{F}, +, \cdot)$ *is a set* $\mathbb{F}$ *endowed with two operations* $+$ *and* $\cdot$ *such that for any* $\lambda, \mu, \nu \in \mathbb{F}$ *one has*

(i) $\lambda + \mu \in \mathbb{F}$ *and* $\lambda \cdot \mu \in \mathbb{F}$, *(internal operations)*

(ii) $(\lambda + \mu) + \nu = \lambda + (\mu + \nu)$ *and* $(\lambda \cdot \mu) \cdot \nu = \lambda \cdot (\mu \cdot \nu)$, *(associativity)*

(iii) $\lambda + \mu = \mu + \lambda$ *and* $\lambda \cdot \mu = \mu \cdot \lambda$, *(commutativity)*

(iv) There exist $0, 1 \in \mathbb{F}$ *such that* $\lambda + 0 = \lambda$ *and* $1 \cdot \lambda = \lambda$, *(existence of identity elements)*

(v) There exists $-\lambda \in \mathbb{F}$ *such that* $\lambda + (-\lambda) = 0$*, and if* $\lambda \neq 0$ *there exists* $\lambda^{-1} \in \mathbb{F}$ *such that* $\lambda \cdot \lambda^{-1} = 1$, *(existence of inverse elements)*

(vi) $\lambda \cdot (\mu + \nu) = \lambda \cdot \mu + \lambda \cdot \nu$. *(distributivity)*

Note that for simplicity, one usually writes $\lambda - \mu$ instead of $\lambda + (-\mu)$ and λ/μ instead of $\lambda \cdot \mu^{-1}$.

Example 3.1.2. *Some examples of fields are* $(\mathbb{R}, +, \cdot)$ *the set of real numbers together with the usual addition and multiplication,* $(\mathbb{Q}, +, \cdot)$ *the set of fractional numbers* $\mathbb{Q} = \big\{a/b \mid a, b \in \mathbb{Z} \text{ with } b \neq 0\big\}$ *together with the usual addition and multiplication,* $(\mathbb{C}, +, \cdot)$ *the set of complex numbers together with its addition and multiplication (as we shall see at the end of this course). Note that for simplicity, one usually writes* $\mathbb{R}, \mathbb{Q}, \mathbb{C}$*, the other two operations being implicit.*

Let us provide a more general framework for the elements of X, Y, Z or $\mathcal{A}, \mathcal{B}, \mathcal{C}$ and for the properties stated in the table above. However, you can always think about $\mathbb{R}^n$ or $M_{mn}(\mathbb{R})$ as the main examples for the following definition. Note that two slightly different fonts are used for the two different multiplications and for the two different additions.

Definition 3.1.3. A vector space over a field $(\mathbb{F}, +, \cdot)$ *consists in a set* V *endowed with two operations* $\boldsymbol{+} : V \times V \to V$ *and* $\bullet : \mathbb{F} \times V \to V$ *such that if* $X, Y, Z \in V$ *and* $\lambda, \mu \in \mathbb{F}$ *the following properties are satisfied:*

(i) $(X \boldsymbol{+} Y) \boldsymbol{+} Z = X \boldsymbol{+} (Y \boldsymbol{+} Z)$,

(ii) $X \boldsymbol{+} Y = Y \boldsymbol{+} X$,

(iii) There exists (a unique) $\mathbf{0} \in V$ *such that* $X \boldsymbol{+} \mathbf{0} = \mathbf{0} \boldsymbol{+} X = X$,

(iv) For all $X \in V$ *there exists* $-X \in V$ *such that* $X \boldsymbol{+} (-X) = \mathbf{0}$,

(v) $\lambda \bullet X \in V$ *and* $1 \bullet X = X$,

(vi) $\lambda \bullet (X + Y) = \lambda \bullet X + \lambda \bullet Y$,

(vii) $(\lambda + \mu) \bullet X = \lambda \bullet X + \mu \bullet X$,

(viii) $(\lambda \cdot \mu) \bullet X = \lambda \bullet (\mu \bullet X)$.

Before providing some examples, let us just mention a consequence of the previous conditions, namely $0 \bullet X = \mathbf{0}$ for any $X \in V$. Indeed, for any $X \in V$ one has

$$X = 1 \bullet X = (1 + 0) \bullet X = 1 \bullet X + 0 \bullet X = X + 0 \bullet X,$$

from which one infers that $0 \bullet X = \mathbf{0}$. Let us also note that whenever the field $\mathbb{F}$ consists in $\mathbb{R}$, one simply says *a real vector space* instead of a vector space over the field $\mathbb{R}$.

Examples 3.1.4. *(i)* $\mathbb{F} = \mathbb{R}$ *and* $V = \mathbb{R}^n$ *with the addition and multiplication by a scalar, as introduced in Chapter 1,*

(ii) $\mathbb{F} = \mathbb{R}$ *and* $V = M_{mn}(\mathbb{R})$ *with the addition and the multiplication by a scalar, as introduced in Chapter 2. More generally, for any field* $\mathbb{F}$ *the set* $M_{mn}(\mathbb{F})$*, defined exactly as* $M_{mn}(\mathbb{R})$*, is a vector space over* $\mathbb{F}$*,*

(iii) $\mathbb{F} = \mathbb{R}$ *and* V *is the set of real functions defined on* $\mathbb{R}$*, with the addition of functions and the multiplication by scalar,*

(iv) $\mathbb{F} = \mathbb{R}$ *and* $V = \{\textit{polynomial functions on } \mathbb{R}\}$

$$= \Big\{ f : \mathbb{R} \to \mathbb{R} \mid f(x) = \sum_{j=0}^{m} a_j x^j \ \textit{with } a_j \in \mathbb{R} \Big\},$$

(v) $\mathbb{F} = \mathbb{R}$ *and* $V = \{f : \mathbb{R} \to \mathbb{R} \textit{ continuous }\}$,

(vi) $\mathbb{F} = \mathbb{R}$ *and* $V = \{f : \mathbb{R} \to \mathbb{R} \textit{ differentiable }\}$.

From now and for simplicity we shall no more use two different notations for the two different multiplications and for the two different additions. This simplification should not lead to any confusion. In addition, we shall simply write $\mathbb{F}$ for the field, instead of $(\mathbb{F}, +, \cdot)$, and the multiplication will be denoted without a dot; the sign $\cdot$ will be kept for the scalar product only.

Definition 3.1.5. *Let* V *be a vector space over* $\mathbb{F}$*, and let* W *be a (non-void) subset of* V*. Then* W *is* a subspace of V *if the following conditions are satisfied:*

(i) If $X, Y \in W$*, then* $X + Y \in W$*,*

(ii) If $X \in W$ *and* $\lambda \in \mathbb{F}$*, then* $\lambda X \in W$*.*

In other words, a subspace of a vector space V is a subset W of V which is stable for the two operations, *i.e.* the addition and the multiplication by a scalar. The next statement will be very useful when checking that a certain set is a vector space. Its proof will be provided in Exercise 3.5.

Lemma 3.1.6. *Any subspace W of a vector space V over a field $\mathbb{F}$ is itself a vector space over $\mathbb{F}$.*

Examples 3.1.7. *(i) $\big\{{}^t(x_1, x_2, \dots, x_{n-1}, 0) \mid x_j \in \mathbb{R}$ for any $j \in \{1, \dots, n-1\}\big\} \subset \mathbb{R}^n$ is a subspace of $\mathbb{R}^n$,*

(ii) If $P, N \in \mathbb{R}^n$ with $N \neq \mathbf{0}$, then $H_{P,N}$ is a subspace of $\mathbb{R}^n$ if and only if $\mathbf{0} \in H_{P,N}$. In particular, $H_{\mathbf{0},N}$ is a subspace of $\mathbb{R}^n$,

(iii) The set of upper triangular $n \times n$ matrices is a subspace of $M_n(\mathbb{R})$,

(iv) The set of $n \times n$ symmetric matrices is a subspace of $M_n(\mathbb{R})$.

The following statement deals with the intersection of subspaces or with the sum of subspaces of a vector space.

Lemma 3.1.8. *Let V be a vector space over a field $\mathbb{F}$, and let W_1, W_2 be two subspaces of V. Then*

(i) $W_1 \cap W_2 = \big\{X \in W_1$ and $X \in W_2\big\}$ is a subspace of V,

(ii) $W_1 + W_2 = \big\{X = X_1 + X_2 \mid X_1 \in W_1$ and $X_2 \in W_2\big\}$ is a subspace of V.

Proof. The proof consists in checking that the two conditions of Definition 3.1.5 are satisfied.

(i) If $X, Y \in W_j$ for $j \in \{1, 2\}$, then $X + Y \in W_j$ because W_j is a subspace. In particular, this implies that if $X, Y \in W_1 \cap W_2$, then $X + Y \in W_1 \cap W_2$. Similarly, in this case one also has $\lambda X \in W_1 \cap W_2$, since W_1 and W_2 are stable for the multiplication by a scalar.

(ii) If $X = X_1 + X_2$ and $Y = Y_1 + Y_2$ with $X_j, Y_j \in W_j$, then $X + Y = X_1 + X_2 + Y_1 + Y_2 = (X_1 + Y_1) + (X_2 + Y_2)$ with $(X_1 + Y_1) \in W_1$ and $(X_2 + Y_2) \in W_2$, which implies that $X + Y \in W_1 + W_2$. Similarly, in this case one also has $\lambda X = \lambda X_1 + \lambda X_2 \in W_1 + W_2$, since both W_1 and W_2 are stable for the multiplication by a scalar. □

3.2 Linear combinations

Let us start with a definition:

Definition 3.2.1. *Let V be a vector space over a field $\mathbb{F}$, and let $X_1, \dots, X_r \in V$. One sets*

$$\mathrm{Vect}(X_1, \dots, X_r) := \big\{\lambda_1 X_1 + \cdots + \lambda_r X_r \mid \lambda_j \in \mathbb{F} \text{ for } j \in \{1, \dots, r\}\big\},$$

and call this set the subspace of V *generated by $X_1, \dots, X_r$.*

Obviously, the first thing to do is to check that this set is indeed a subspace of V.

Lemma 3.2.2. *In the above setting,* $\mathrm{Vect}(X_1,\dots,X_r)$ *is a subspace of* V*.*

Proof. The proof consists in checking that both conditions of Definition 3.1.5 are satisfied. First of all, if $X = \lambda_1 X_1 + \dots + \lambda_r X_r$ and $X' = \lambda_1' X_1 + \dots + \lambda_r' X_r$, then

$$X + X' = \underbrace{(\lambda_1 + \lambda_1')}_{\in \mathbb{F}} X_1 + \dots + \underbrace{(\lambda_r + \lambda_r')}_{\in \mathbb{F}} X_r \ \in \mathrm{Vect}(X_1,\dots,X_r).$$

Similarly, if $X = \lambda_1 X_1 + \dots + \lambda_r X_r$ and $\lambda \in \mathbb{F}$, then

$$\lambda X = \underbrace{(\lambda\lambda_1)}_{\in \mathbb{F}} X_1 + \dots + \underbrace{(\lambda\lambda_r)}_{\in \mathbb{F}} X_r \ \in \mathrm{Vect}(X_1,\dots,X_r).$$

Since both conditions are checked, it is thus a subspace of V. □

Since $\mathrm{Vect}(X_1,\dots,X_r)$ is a subspace of V, it was legitimate to call it as we did. Note that one also says that $\lambda_1,\dots,\lambda_r$ are *the coefficients* of *the linear combination* $\lambda_1 X_1 + \dots + \lambda_r X_r$.

Remark 3.2.3. *If* $\mathrm{Vect}(X_1,\dots,X_r) = V$*, then one says that* V *is generated by* $X_1,\dots,X_r$*, or that* $\{X_1,\dots,X_r\}$ *is a* generating family*.*

The following three examples are related to real vector spaces, as it is the case in most of the examples of these lecture notes.

Examples 3.2.4. *(i) Recall that* $E_j = {}^t(0,\dots,1,\dots,0)$ *with the entry* 1 *at the position* j*. Then* $\{E_j\}_{j=1}^n \equiv \{E_1, E_2, \dots, E_n\}$ *is a generating family for* $\mathbb{R}^n$*.*

(ii) If $N \in \mathbb{R}^n$ *with* $N \neq \mathbf{0}$*, then* $\mathrm{Vect}(N)$ *is the line passing through* $\mathbf{0}$ *and having the direction* N*,* i.e. $\mathrm{Vect}(N) = L_{\mathbf{0},N}$*, with the* $L_{\mathbf{0},N}$ *defined in Definition 1.5.1.*

(iii) If $X, Y \in \mathbb{R}^3$ *with* $X \neq \mathbf{0}$*,* $Y \neq \mathbf{0}$*, and* $Y \neq \lambda X$ *for any* $\lambda \in \mathbb{R}$*, then* $\mathrm{Vect}(X,Y)$ *defines a plane in* $\mathbb{R}^3$ *passing through* $\mathbf{0}$*. In fact, it corresponds to the plane passing through the three points* $\mathbf{0}, X, Y$*, as seen in Exercise 1.18.*

Remark 3.2.5. *If* $\mathbb{F} = \mathbb{R}$ *and if one considers* $X_1,\dots,X_r \in \mathbb{R}^n$*, then one can set*

$$\mathrm{Box}(X_1,\dots,X_r) := \big\{\lambda_1 X_1 + \dots + \lambda_r X_r \mid \lambda_j \in [0,1] \text{ for } j \in \{1,\dots,r\}\big\}.$$

This is a subset of $\mathrm{Vect}(X_1,\dots,X_r)$*, called* the hyperbox *or* generalized box *generated by* $X_1,\dots X_r$*. Note that* $\mathrm{Box}(X_1,\dots,X_r)$ *is not a subspace. It is also easily observed that* $\mathrm{Box}(X_1)$ *corresponds to the segment between* $\mathbf{0}$ *and* X_1 *and that* $\mathrm{Box}(X_1,X_2)$ *corresponds to the parallelogram generated by* X_1 *and* X_2*, and with one apex at* $\mathbf{0}$*.*

3.3 Convex sets

In this section we consider only real vector spaces, *i.e.* the field $\mathbb{F}$ is equal to $\mathbb{R}$ for all vector spaces.

Definition 3.3.1. *Let S be a subset of a real vector space V. Then S is* convex *if for any $X, Y \in S$ and for any $t \in [0,1]$ one has*

$$X + t(Y - X) \equiv (1-t)X + tY \ \in S.$$

Examples 3.3.2. *(i) A ball is convex, but a doughnut is not convex,*

(ii) For any $X_1, \dots, X_r$ in a real vector space V, $\mathrm{Vect}(X_1, \dots, X_r)$ is convex. Indeed, if $X = \lambda_1 X_1 + \dots + \lambda_r X_r$ and $X' = \lambda'_1 X_1 + \dots + \lambda'_r X_r$ with $\lambda_j, \lambda'_j \in \mathbb{R}$ then

$$(1-t)X + tX' = \underbrace{\big((1-t)\lambda_1 + t\lambda'_1\big)}_{\in \mathbb{R}} X_1 + \underbrace{\big((1-t)\lambda_r + t\lambda'_r\big)}_{\in \mathbb{R}} X_r \ \in \mathrm{Vect}(X_1, \dots, X_r),$$

(iii) For any $X_1, \dots, X_r$ in a real vector space V, $\mathrm{Box}(X_1, \dots, X_r)$ is convex. Indeed, in the framework of the previous example, observe that if $0 \leq \lambda_j \leq 1$ and $0 \leq \lambda'_j \leq 1$ then one has for any $t \in [0,1]$

$$0 \leq (1-t)\lambda_j + t\lambda'_j \leq (1-t)1 + t1 = 1.$$

As a consequence, one infers that

$$(1-t)X + tX' = \underbrace{\big((1-t)\lambda_1 + t\lambda'_1\big)}_{\in [0,1]} X_1 + \underbrace{\big((1-t)\lambda_r + t\lambda'_r\big)}_{\in [0,1]} X_r \ \in \mathrm{Box}(X_1, \dots, X_r).$$

Definition 3.3.3. *Let V be a real vector space, and let $X_1, \dots, X_r \in V$. We set*

$$\mathrm{CS}(X_1, \dots, X_r) := \big\{\lambda_1 X_1 + \dots + \lambda_r X_r \mid 0 \leq \lambda_j \leq 1 \text{ for } j \in \{1, \dots, r\} \\ \text{and } \lambda_1 + \lambda_2 + \dots + \lambda_r = 1\big\}$$

and call is the convex set *generated or spanned by $X_1, \dots, X_r$.*

Note that by definition, the following inclusions always hold

$$\mathrm{CS}(X_1, \dots, X_r) \subset \mathrm{Box}(X_1, \dots, X_r) \subset \mathrm{Vect}(X_1, \dots, X_r).$$

Example 3.3.4. *If $V = \mathbb{R}^n$, then $\mathrm{CS}(X_1)$ corresponds just to the point X_1, $\mathrm{CS}(X_1, X_2)$ corresponds to the segment between X_1 and X_2 while $\mathrm{CS}(X_1, X_2, X_3)$ corresponds to the triangle of apexes X_1, X_2 and X_3.*

Obviously, one has to show immediately the following statement:

Lemma 3.3.5. *If V is a real vector space and $X_1, \ldots, X_r \in V$, then* $\mathrm{CS}(X_1, \ldots, X_r)$ *is convex.*

In fact, $\mathrm{CS}(X_1, \ldots, X_r)$ is the smallest convex set containing $X_1, \ldots, X_r$.

Proof. Let $X = \lambda_1 X_1 + \cdots + \lambda_r X_r$ with $0 \leq \lambda_j \leq 1$ and $\lambda_1 + \lambda_2 + \cdots + \lambda_r = 1$, and let $X' = \lambda'_1 X_1 + \cdots + \lambda'_r X_r$ with $0 \leq \lambda'_j \leq 1$ and $\lambda'_1 + \lambda'_2 + \cdots + \lambda'_r = 1$. Then for any $t \in [0,1]$ one has

$$(1-t)X + tX' = \sum_{j=1}^{r} \big((1-t)\lambda_j + t\lambda'_j\big) X_j$$

with $0 \leq (1-t)\lambda_j + t\lambda'_j \leq (1-t)1 + t1 = 1$ and

$$\sum_{j=1}^{r} \big((1-t)\lambda_j + t\lambda'_j\big) = (1-t)\sum_{j=1}^{r} \lambda_j + t\sum_{j=1}^{r} \lambda'_j = (1-t)1 + t1 = 1.$$

As a consequence, $(1-t)X + tX' \in \mathrm{CS}(X_1, \ldots, X_r)$, which means that $\mathrm{CS}(X_1, \ldots, X_r)$ is convex. □

3.4 Linear independence

The following definition will be of importance in the sequel.

Definition 3.4.1. *Let V be a vector space over a field $\mathbb{F}$, and let $X_1, \ldots, X_r \in V$. The elements $X_1, \ldots, X_r$ are* linearly dependent *if there exist $\lambda_1, \ldots, \lambda_r \in \mathbb{F}$ not all equal to 0 such that*

$$\lambda_1 X_1 + \cdots + \lambda_r X_r = \mathbf{0}. \tag{3.4.1}$$

The elements $X_1, \ldots, X_r$ are said linearly independent *if there do not exist such scalars $\lambda_1, \ldots, \lambda_r$.*

Note that alternatively, the vectors $X_1, \ldots, X_r$ are linearly independent if whenever (3.4.1) is satisfied, then one must have $\lambda_1 = \lambda_2 = \cdots = \lambda_r = 0$. In this case, one also says that the family $\{X_1, \ldots, X_r\}$ is linearly independent.

Examples 3.4.2. *(i) For $V = \mathbb{R}^n$, the family $\{E_j\}_{j=1}^n$ is linearly independent,*

(ii) For $V = M_n(\mathbb{R})$, the family $\{I_{rs}\}_{r,s=1}^n$ of elementary matrices introduced in Section 2.5 is linearly independent,

(iii) For $V = \mathbb{R}^2$, the elements $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$, and $\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$ are linearly dependent since

$$1\begin{pmatrix}1\\0\end{pmatrix} + 1\begin{pmatrix}0\\1\end{pmatrix} - 1\begin{pmatrix}1\\1\end{pmatrix} = \mathbf{0}.$$

(iv) *Let f_1, f_2 be two continuous real functions on $\mathbb{R}$. In this case f_1, f_2 are linearly dependent if there exists $\lambda_1, \lambda_2 \in \mathbb{R}$ with $(\lambda_1, \lambda_2) \neq (0,0)$ such that $\lambda_1 f_1 + \lambda_2 f_2 = \mathbf{0}$, or more precisely*

$$\lambda_1 f_1(x) + \lambda_2 f_2(x) = 0 \quad \text{for all } x \in \mathbb{R}.$$

For example, if $f_1(x) = \cos(x)$ and $f_2(x) = \sin(x)$, then f_1 and f_2 are linearly independent even if $0\cos(0) + \lambda \sin(0) = 0$ for arbitrary $\lambda \in \mathbb{R}$.

Definition 3.4.3. *Let V be a vector space over a field $\mathbb{F}$, and let $X_1, \dots, X_r \in V$. If $\mathrm{Vect}(X_1, \dots, X_r) = V$ and if $X_1, \dots, X_r$ are linearly independent, then $\{X_1, \dots, X_r\}$ is called* a basis *for V. Alternatively, one also says that the family $\{X_1, \dots, X_r\}$ constitutes or forms a basis for V.*

Examples 3.4.4. (i) *$\{E_j\}_{j=1}^n$ forms a basis for $\mathbb{R}^n$,*

(ii) *$\{I_{rs}\}_{r,s=1}^n$ forms a basis for $M_n(\mathbb{R})$,*

(iii) *$\{x \mapsto x^n\}_{n=0}^\infty$ forms a basis for the vector space of all polynomials on $\mathbb{R}$.*

Let us consider a special case of the previous definition in the case $n = 2$. The content of the following lemma will be useful later on, and its proof will be provided in Exercise 3.9.

Lemma 3.4.5. *Let $\binom{a}{b}$, $\binom{c}{d} \in \mathbb{R}^2$, with $a, b, c, d \in \mathbb{R}$.*

(i) *The two vectors are linearly independent if and only if $ad - bc \neq 0$,*

(ii) *If the two vectors are linearly independent, they form a basis of $\mathbb{R}^2$.*

Given a basis of a vector space V, any point X can be expressed as a linear combinations of elements of this basis. More precisely, one sets:

Definition 3.4.6. *Let $\{X_1, \dots, X_r\}$ be a basis for a vector space V over $\mathbb{F}$. Then, for $X = \lambda_1 X_1 + \dots + \lambda_r X_r$ the coefficients $\{\lambda_1, \dots, \lambda_r\}$ are called* the coordinates of X with respect to the basis $\{X_1, \dots, X_r\}$ of V.

In order to speak about "the" coordinates, the following lemma is necessary.

Lemma 3.4.7. *The coordinates of a vector with respect to a basis are unique.*

Proof. Let $\{X_1, \dots, X_r\}$ be a basis, and assume that

$$X = \lambda_1 X_1 + \dots + \lambda_r X_r = \lambda_1' X_1 + \dots + \lambda_r' X_r.$$

It then follows that

$$X - X = \mathbf{0} = (\lambda_1 - \lambda_1')X_1 + \dots + (\lambda_r - \lambda_r')X_r.$$

By independence of $X_1, \dots, X_r$, it follows that $(\lambda_j - \lambda_j') = 0$ for all $j \in \{1, \dots, r\}$, which means that $\lambda_j = \lambda_j'$. Thus, the coordinates of X with respect to a basis are unique. □

3.5 Dimension

Question: Can one find 3 linearly independent elements in $\mathbb{R}^2$? For instance, if $A = \binom{1}{2}$, $B = \binom{-5}{7}$ and $C = \binom{10}{4}$, are they linearly independent vectors ? The answer is no, and there is no need to do any computation for getting this answer. Indeed, let us consider the more general setting provided by $A = \binom{a_1}{a_2}$, $B = \binom{b_1}{b_2}$ and $C = \binom{c_1}{c_2}$. Then one has

$$\lambda_1 \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \lambda_2 \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + \lambda_3 \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \mathbf{0} \Longleftrightarrow \left\{ \begin{array}{l} \lambda_1 a_1 + \lambda_2 b_1 + \lambda_3 c_1 = 0 \\ \lambda_1 a_2 + \lambda_2 b_2 + \lambda_3 c_2 = 0 \end{array} \right. .$$

Note that this corresponds to a system of two equations for the three unknowns λ_1, λ_2 and λ_3. As seen in Theorem 2.3.4, such a homogeneous system of equations has always a non-trivial solution, which means that there exists a non trivial solution for the corresponding equation 3.4.1. As a consequence, three vectors in $\mathbb{R}^2$ can never be independent.

More generally, one has:

Theorem 3.5.1. *Let $\{X_1, \dots, X_r\}$ be a basis of a vector space V over $\mathbb{F}$. Let $Y_1, \dots, Y_m \in V$ and assume that $m > r$. Then $Y_1, \dots, Y_m$ are linearly dependent.*

Note that if $m \leq r$, the statement does not imply that $Y_1, \dots, Y_m$ are linearly independent.

Now, in order to give the proof in its full generality, we need to extend the definition of $M_{mn}(\mathbb{R})$ to $M_{mn}(\mathbb{F})$, for an arbitrary field $\mathbb{F}$. In fact, since elements in a field can be added and multiplied, all definitions related to $M_{mn}(\mathbb{R})$ can be translated directly into the same definitions for $M_{mn}(\mathbb{F})$. The only modification is that any entry a_{ij} of a matrix $\mathcal{A}$ belongs to $\mathbb{F}$ instead of $\mathbb{R}$, and the multiplication of a matrix by a scalar $\lambda \in \mathbb{R}$ is now replaced by the multiplication by an element of $\mathbb{F}$. Then, most of the statements of Section 2 are valid (simply by replacing $\mathbb{R}$ by $\mathbb{F}$), and in particular Theorem 2.3.4 can be obtained in this more general context. This theorem is precisely the one required for the proof of the above statement.

Proof. Since $X_1, \dots, X_r$ generate V, there exists $a_{ij} \in \mathbb{F}$ such that

$$Y_j = a_{1j} X_1 + \cdots + a_{rj} X_r \quad \text{for any } j \in \{1, \dots, m\}.$$

Then, let us consider $\lambda_1, \dots, \lambda_m \in \mathbb{F}$ and observe that

$$\begin{aligned} & \lambda_1 Y_1 + \cdots + \lambda_m Y_m = \mathbf{0} \\ \Longleftrightarrow & \lambda_1 (a_{11} X_1 + \cdots + a_{r1} X_r) + \cdots + \lambda_m (a_{1m} X_1 + \cdots + a_{rm} X_r) = \mathbf{0} \\ \Longleftrightarrow & (\lambda_1 a_{11} + \lambda_2 a_{12} + \cdots + \lambda_m a_{1m}) X_1 + \cdots + (\lambda_1 a_{r1} + \dots \lambda_m a_{rm}) X_r = \mathbf{0} \end{aligned}$$

which is still equivalent to the following expression, by linear independence of $X_1, \dots, X_r$:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \dots & a_{rm} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Finally, since $m > r$, it follows from a simple adaptation of Theorem 2.3.4 to $M_n(\mathbb{F})$ that this system of equation always has a non-trivial solution. However, this means precisely that the elements $Y_1, \dots, Y_m$ are linearly dependent. □

Corollary 3.5.2. *Let V be a vector space and suppose that $\{X_1, \dots, X_n\}$ is a basis for V. Then any other basis for V also contains n elements.*

Proof. Let $\{Y_1, \dots, Y_m\}$ be a second basis for V. If $m > n$ then $Y_1, \dots, Y_m$ can not be linearly independent, by the previous theorem. Similarly, if $m < n$ then $X_1, \dots, X_n$ can not be linearly independent, also by the previous theorem. Since any basis is made of linearly independent vectors, one obtains that the only possibility is $m = n$. □

We now define a notion which has been implicitly used from the beginning for $\mathbb{R}^n$.

Definition 3.5.3. *Let V be a vector space over a field $\mathbb{F}$, and let $\{X_1, \dots, X_n\}$ be a basis for V. Then one says that V is* of dimension n*, since this number is independent of the choice of a particular basis for V. The dimension of the vector space V is denoted by* $\dim(V)$*.*

Remark 3.5.4. *In these lecture notes except in a few exercises, all vector spaces are of finite dimension. This fact is tacitly assumed in many statements later on, but note that vector spaces of infinite dimensions often appear in physics or in mathematics.*

Examples 3.5.5. *(i) $\mathbb{R}^n$ is of dimension n, $\mathbb{F}^n$ is also of dimension n,*

(ii) For any vector space V and any $X \in V$, the dimension of $\mathrm{Vect}(X)$ *is* 1*,*

(iii) Any plane in $\mathbb{R}^3$ (passing through the origin) is of dimension 2*, while a line passing through the origin is of dimension* 1*.*

The following result is often useful, when the dimension of the vector space is already known.

Lemma 3.5.6. *Let V be a vector space of dimension n, and let $X_1, \dots, X_n \in V$ be linearly independent. Then $\{X_1, \dots, X_n\}$ is a basis for V.*

Proof. One only has to show that $\mathrm{Vect}(X_1, \dots, X_n) = V$. By contradiction, assume that there exists $Y \in V$ such that $Y \notin \mathrm{Vect}(X_1, \dots, X_n)$. Then, the vectors $X_1, \dots, X_n, Y$ are linearly independent, and $\{X_1, \dots, X_n, Y\}$ would generate a basis for V of dimension $n + 1$, which is impossible by the previous Corollary. □

3.6 The rank of a matrix

Let $\mathcal{A} \in M_{mn}(\mathbb{F})$, and recall from Section 2.2 that the columns of $\mathcal{A}$ have been denoted by $\mathcal{A}^1, \dots, \mathcal{A}^n$. Each column is an element of $\mathbb{F}^m$ (a column vector with m entries in $\mathbb{F}$) and the family $\{\mathcal{A}^1, \dots, \mathcal{A}^n\}$ generates a subspace of $\mathbb{F}^m$, which we have denoted by

$\mathrm{Vect}(\mathcal{A}^1, \ldots, \mathcal{A}^n)$. In this case this subspace is called *the column space*. Alternatively, the rows of $\mathcal{A}$ generate the subspace $\mathrm{Vect}({}^t\mathcal{A}_1, \ldots, {}^t\mathcal{A}_m)$ of $\mathbb{F}^n$, which is called *the row space*. The dimension of the first subspace is called *the column rank*, while the dimension of the second subspace is called *the row rank*.

By what we have seen in the previous sections, the column rank corresponds to the maximal number of linearly independent columns, while the row rank corresponds to the maximal number of linearly independent rows. Our aim in this section is to study these numbers.

Lemma 3.6.1. *Elementary row operations do not change the row rank of a matrix.*

Proof. For this proof, it is sufficient to observe that

$$\begin{aligned}
&\mathrm{Vect}({}^t\mathcal{A}_1, \ldots, {}^t\mathcal{A}_j, \ldots, {}^t\mathcal{A}_k, \ldots, {}^t\mathcal{A}_m) \\
&= \mathrm{Vect}({}^t\mathcal{A}_1, \ldots, {}^t\mathcal{A}_k, \ldots, {}^t\mathcal{A}_j, \ldots, {}^t\mathcal{A}_m) \\
&= \mathrm{Vect}({}^t\mathcal{A}_1, \ldots, c{}^t\mathcal{A}_j, \ldots, {}^t\mathcal{A}_k, \ldots, {}^t\mathcal{A}_m) \\
&= \mathrm{Vect}({}^t\mathcal{A}_1, \ldots, {}^t\mathcal{A}_j + c{}^t\mathcal{A}_k, \ldots, {}^t\mathcal{A}_k, \ldots, {}^t\mathcal{A}_m)
\end{aligned}$$

for any $c \neq 0$. Since these subspaces are the same, their dimension coincide. □

Similarly, one can show that elementary row operations do not change the column rank of a matrix. Note that this proof is less easy since elementary row operations have been defined on rows, and not on columns.

Theorem 3.6.2. *For any $\mathcal{A} \in M_{mn}(\mathbb{F})$, the column rank and the row rank of $\mathcal{A}$ are equal.*

Thanks to this statement, it is sufficient to speak about *the rank of a matrix*, denoted by $\mathrm{rank}(\mathcal{A})$, there is no need to specify if it is the column rank or the row rank.

Proof. First of all, recall that the matrix $\mathcal{A}$ is row equivalent to a matrix $\mathcal{B}$ in the standard form, see Corollary 2.4.10. Since elementary row operations do not change the row rank or the column rank, the matrix $\mathcal{B}$ has the same row rank and column rank as the original matrix $\mathcal{A}$. Then, it is easily observed that the number of leading coefficients of $\mathcal{B}$ is equal to the number of linearly independent rows, but also to the number of linearly independent columns of $\mathcal{B}$. Therefore, the number of leading coefficients of $\mathcal{B}$ is equal to the row rank of $\mathcal{B}$ and to the column rank of $\mathcal{B}$. It follows that these two numbers are equal, and that the row rank of $\mathcal{A}$ and the column rank of $\mathcal{A}$ are also equal to this number. □

3.7 Exercises

Exercise 3.1. *Show that the following sets of elements of* $\mathbb{R}^3$ *form subspaces :*

i) $S_1 := \{{}^t(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$,

ii) $S_2 := \{{}^t(x, y, z) \in \mathbb{R}^3 \mid x = y \text{ and } 2y = z\}$,

iii) $S_3 := \{{}^t(x, y, z) \in \mathbb{R}^3 \mid x + y = 3z\}$.

Exercise 3.2. *Let* V *be a subspace of* $\mathbb{R}^n$*, and let* W *be the set of all elements of* $\mathbb{R}^n$ *which are perpendicular to all elements of* V*. Show that* W *itself is a subspace of* $\mathbb{R}^n$*. This subspace is often denoted by* $V^\perp$ *and called the orthogonal complement of* V *in* $\mathbb{R}^n$*.*

Exercise 3.3. *Let* $A_1, \dots, A_r$ *be generators of a subspace* V *of* $\mathbb{R}^n$*. Let* W *be the set of all elements in* $\mathbb{R}^n$ *which are perpendicular to* $A_1, \dots, A_r$*. Show that* $W = V^\perp$*.*

Exercise 3.4. *Show that the set of all real polynomials is a subspace of the vector space of all real and continuous functions on* $\mathbb{R}$*. Exhibit a generating family for this subspace.*

Exercise 3.5. *Let* V *be a vector space over a field* $\mathbb{F}$*. Show that any subspace of* V *is itself a vector space.*

Exercise 3.6. *Let* S *be a convex set in a real vector space* V*.*

i) For $\lambda \in \mathbb{R}$*, show that* λS *is a convex set in* V*, with* $\lambda S = \{\lambda X \mid X \in S\}$*.*

ii) For $Y \in V$*, show that* $S + Y$ *is a convex set in* V*, with* $S + Y = \{X + Y \mid X \in S\}$*.*

Exercise 3.7. *Show that the intersection of two convex sets is still convex.*

Exercise 3.8. *Show that the vectors* $\begin{pmatrix}1\\0\\0\end{pmatrix}$, $\begin{pmatrix}0\\1\\0\end{pmatrix}$ *and* $\begin{pmatrix}1\\1\\1\end{pmatrix}$ *form a basis of* $\mathbb{R}^3$*.*

Exercise 3.9. *Let* $\begin{pmatrix}a\\b\end{pmatrix}, \begin{pmatrix}c\\d\end{pmatrix} \in \mathbb{R}^2$*. Show that these two vectors are linearly independent if and only if* $ad - bc \neq 0$*.*

Exercise 3.10. *Express the coordinates of* Y *in the basis generated by* X_1 *and* X_2 *:*

i) $Y = \begin{pmatrix}1\\0\end{pmatrix}$, $X_1 = \begin{pmatrix}1\\1\end{pmatrix}$ *and* $X_2 = \begin{pmatrix}0\\1\end{pmatrix}$,

ii) $Y = \begin{pmatrix}2\\1\end{pmatrix}$, $X_1 = \begin{pmatrix}1\\-1\end{pmatrix}$ *and* $X_2 = \begin{pmatrix}1\\1\end{pmatrix}$.

Exercise 3.11. *Let* $X_1, \dots, X_r$ *be non-zero elements of* $\mathbb{R}^n$ *and assume that* $X_j \cdot X_k = 0$ *for each* $j \neq k$*. Show that these elements are linearly independent.*

Exercise 3.12. *Determine the dimension of the following subspaces:*

i) $S_1 := \{{}^t(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$, *ii)* $S_2 := \{{}^t(x, y, z) \in \mathbb{R}^3 \mid x = y \text{ and } 2y = z\}$,

iii) $S_3 := \{{}^t(x, y, z) \in \mathbb{R}^3 \mid x + y = 3z\}$.

Exercise 3.13. *Determine the rank of the following matrices :*

$$a) \begin{pmatrix} 6 & 3 & -4 \\ -4 & 1 & -6 \\ 1 & 2 & -5 \end{pmatrix} \quad b) \begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{pmatrix} \quad c) \begin{pmatrix} 0 & 1 & 3 & -2 \\ 2 & 1 & -4 & 3 \\ 2 & 3 & 2 & -1 \end{pmatrix} \quad d) \begin{pmatrix} 1 & 2 & -1 & 2 & 1 \\ 2 & 4 & 1 & -2 & 3 \\ 3 & 6 & 2 & -6 & 5 \end{pmatrix}$$

Exercise 3.14. *A* doubly stochastic matrix *is a $n \times n$ matrix $\mathcal{A} = (a_{jk})$ such that $a_{jk} \in [0, 1]$ and such that the sum of the elements of each line is equal to 1, as well as the sum of the elements of each column.*

(i) Show that the product of two doubly stochastic matrices is still a doubly stochastic matrix,

(ii) Show that the set of all doubly stochastic matrices is a convex set.